Method of Burst Scheduling in a Communication Network

Field of the Invention

This invention relates to the scheduling of burst transmissions in communication networks.

Art Background

The routing of transmissions through optical networks of the prior art is in many cases performed actively, using thermo-optical or optomechanical switches, or using devices that require conversion from the optical to the electrical domain and back. Such active devices tend to add significant expense to the network infrastructure, and generally dissipate significant amounts of power.

One alternative approach to network architecture that avoids the above disadvantages is Time-Domain Wavelength-Interleaved Network (TWIN). TWIN is transparent, in the sense that each transmitted signal remains in the optical domain, at the same wavelength, all the way from the source node to the destination node. A unique wavelength (or set of wavelengths) is assigned to each destination node. Each source node is served by a multiwavelength source. Outgoing transmissions are packaged in the form of optical bursts, each in a specified wavelength channel. In traversing the network, each signal passes through one or more wavelength-selective cross-connects, which ultimately steer each burst to its intended destination. Thus, the wavelength channel assigned to an outgoing burst is, in effect, an intrinsic destination label for that burst.

In the absence of coordination among the source nodes, it is possible for bursts to collide at the destination. That is, bursts from two disparate sources might wholly or partially coincide when received at the destination node. The resulting interference will often be irresolvable. It is also possible in some cases for a single source to be scheduled to send out bursts to two disparate destinations in the same timeslot. Although an occurrence of this kind is also in some sense a "collision," we have here chosen to refer to it as a "source conflict." One solution to the problem of collisions, as well as of source conflicts, has been described in U.S. Patent Application Serial No. 10/426,389 which was filed by K. Kumaran et al. on April 30, 2003 and is commonly assigned herewith. The solution described there involves a centralized

scheduler which coordinates the source nodes, while also taking into account the respective transmission delays between each source-destination pair.

Although coordinated scheduling provides a useful solution to the problem of collisions, there are also potential advantages to alternative solutions, particularly

5   when traffic is asynchronous, i.e., characterized by dynamically varying bandwidth demands.

Summary of the Invention

We have discovered that a TWIN network subject to distributed scheduling can support useful levels of throughput, even without coordination among the source

10   nodes. Accordingly, our invention involves a method for scheduling a burst transmission in a network for delivery to a destination. The network may be an optical network. However, the invention also has applications to other communication media. For example, the invention also may be applied in the context of wireless local area networks and other types of radiofrequency communication

15   network.

According to our method, in one aspect, at a receiving node of a communication network, a request is received to schedule at least one timeslot of a recurrent cycle for receipt of burst transmissions from a sending node of the network. In response to the scheduling request, at least one timeslot of the cycle is selected in a

20   manner which is independent of timeslot selections made by other nodes of the network for receipt of burst transmissions. The selected timeslot or timeslots are communicated to the sending node.

In specific embodiments of the invention, each destination node makes timeslot assignments in such a way that no two sources are scheduled for burst

25   arrivals at the destination within the same timeslot. Thus, collisions between arriving bursts at the destination are avoided.

Specific embodiments of the invention include an adaptive feature that tends to reduce the number of source conflicts at the source node. According to such feature, if the destination node senses a missed transmission in an assigned timeslot, it

30   will substitute a different timeslot and notify the source node accordingly.

Brief Description of the Drawing

FIG. 1 is a simplified, schematic diagram of an illustrative TWIN network with distributed scheduling.

FIG. 2 is a schematic diagram of a burst cycle useful in the practice of the invention.

FIG. 3 is a modified view of the network of FIG. 1, illustrating a transaction between two nodes of the network.

FIG. 4 is a diagram of an exemplary transaction between a source node and a destination node according to a protocol for distributed scheduling.

FIG. 5 is a state diagram, illustrating the various possible states of a timeslot according to a protocol for distributed scheduling, and of the possible transitions between the states.

FIG. 6 is a graph, based on numerical simulations, which illustrates the blocking performance of a network implementing certain principles of the present invention in an exemplary embodiment.

Detailed Description

Below, we will describe an illustrative embodiment of the invention for use in a TWIN network. It should be noted that procedures and protocols to be described below also have application in communication media of other kinds, including in radiofrequency communication networks. The principles to be set out below are readily transferred to other such contexts. Accordingly, it should be understood that the context of an optical network in which the invention is to be described is merely illustrative, and not limiting as to the scope of the invention.

In the illustrative TWIN network of FIG. 1, nodes 10 are interconnected by optical fibers or other waveguiding media 20, and wavelength-selective cross connects (WSXCs) 30. Each node able to function as a destination for optical bursts is equipped with a distributed scheduler 40. The WSXCs 30 and their interconnecting links constitute the core of the network, and the nodes 10 constitute its edge. In typical applications, each node 10 will serve as an aggregation device which provides service interfaces to clients communication devices situated outside of the TWIN network. The client devices are potentially of various kinds, including IP routers, ATM switches, and SONET ADMs. At the node, incoming client protocol units are

aggregated, encapsulated into optical bursts, and launched into the TWIN network via a fast-tunable laser. Also at the node, outgoing bursts are decapsulated into client protocol units and delivered to the appropriate clients. More details concerning TWIN network architecture can be found in U.S. Patent Application Serial No.

5   10/426,388, filed on April 30, 2003 by R.C. Giles et al. and commonly assigned herewith.

Optical bursts from the source nodes are scheduled to be transmitted within assigned timeslots, as will be explained below. Timeslots occur in recurring cycles, as illustrated in FIG. 2. Turning to the figure, it will be seen that one cycle is divided

10   into B timeslots, numbered 0, 1, ..., B-1. Each timeslot is of duration T_s, including an overhead portion of duration T_0, a payload portion of duration T_p, and a guard band, separating it from adjacent timeslot, of duration T_g.

Importantly, both synchronous and asynchronous traffic can be scheduled and transported on the TWIN network. For synchronous traffic, a constant service rate

15   will typically be provided during the lifetime of a connection. For asynchronous traffic, such as IP traffic, the service rate can be changed dynamically, during the lifetime of a connection, in response to requests issued by the source node. The service rate is measured by the number of timeslots provided per cycle.

Optionally, the cycle can be subdivided into a "synch period" for transmission

20   of synchronous traffic, and an "asynch period" for transmission of asynchronous traffic, as shown in FIG. 2. The boundary between the synch period and the asynch period will typically be adjustable according to the current traffic characteristics.

In FIG. 3, a particular pair of nodes 50S and 50D have been designated, for purposes of illustration, as a respective source and destination node. To initiate a

25   transmission, source node 50S will send a timeslot request to destination node 50D. The source node will request one or more recurring timeslots, which are to be provided in each cycle until a new request is sent. Optionally, the source node may add timeslot requests and other control information to the overhead portions of already-scheduled, outgoing optical bursts, if such bursts are available and intended

30   for the appropriate destination node. More typically, a separate Data Communication Network (DCN) which overlies the TWIN network will be used to communicate

control information between source and destination nodes. A portion of the DCN is represented in FIG. 3 by the broken line 60.

Destination node 50D will respond to a timeslot request from node 50S with a timeslot assignment message. The timeslot assignment will typically be

5   communicated over the DCN, but as noted, may optionally be added to the overhead portion of an appropriate burst. Typically, the information sent back to the source node will identify timeslots for <u>arrival</u> of the bursts at the destination node. Therefore, it will be necessary for the source node to account for transmission delay in the network, and to transmit the burst at a time earlier, by an appropriate amount,

10 than the assigned timeslot. More specifically, a burst assigned the timeslot t_a will depart node 50S at time $t_a - \delta + kB$, wherein δ signifies the transmission delay, and kB signifies the smallest feasible number of whole cycle times. Synchronization of all of the nodes to a common clock can be achieved using a GPS timing reference. Once this is done, the delays δ are readily measurable.

15 As noted above, each source node is equipped with a distributed scheduler. In FIG. 3, for example, node 50S is shown in communication with distributed scheduler 70. The distributed schedulers do not communicate with each other. Typically, each distributed scheduler receives traffic information only in the timeslot requests issued to it by its requesting source nodes. One function of the distributed scheduler is to

20 schedule timeslots. In particular embodiments of the invention, the distributed scheduler performs the scheduling in such a way that collisions at the destination are avoided, and provides substitute timeslots when conflicts occur at a source node. Communication of schedules between the source and destination nodes takes place via the "request" messages from the source, and "grant" messages from the

25 destination.

In principle, the distributed scheduler can be operated in a mode that makes random selections of timeslots for assignment, irrespective of whether any have already been assigned. However, except when the traffic loads are very light, the high blocking probability that results from such a scheme makes it unattractive. Therefore,

30 in the protocol to be described below, the distributed scheduler avoids assigning timeslots that will collide at the destination.

With reference to FIG. 4, we will now describe a protocol for distributed scheduling. The source node together with its distributed scheduler will be referred to below as "Source," and the destination node as "Destination." In message R1, Source requests to schedule one or more burst transmissions per cycle. In the specific

5   example illustrated, Source is requesting a rate of three bursts per cycle. Upon receiving request R1, Destination grants the requested number of timeslots, for use in cycles subsequent to the current cycle. Thus, for example, grant G1 as illustrated in the figure offers the source timeslots 1, 6, and 9. Upon receiving the grant, Source checks the availability of its tunable laser for transmissions in all of the assigned

10   timeslots. In the example illustrated, Source discovers that there is a conflict with another transmission already scheduled for timeslot 6. Consequently, Source transmits only in timeslots 1 and 9, as indicated by edge T1 in the figure.

Destination learns about conflicts by detecting the <u>absence</u> of received bursts on assigned timeslots. Accordingly, in the present example, Destination fails to

15   receive a burst in timeslot 6 on transmission T1, and thereby detects that timeslot 6 is subject to a conflict. Upon detection of a conflict, Destination returns the unused timeslot to the pool of available timeslots, and assigns a new timeslot to Source. In the present example, timeslot 8 is assigned in substitution for timeslot 6. This substitution is communicated to Source in grant G2. In the exemplary scenario shown

20   in the figure, grant G2 arrives at Source during the next cycle following the cycle in which G2 was sent out. In the meantime, Source has sent out a second transmission on timeslots 1 and 9 only, as indicated by edge T2 in the figure. In the next cycle, Source sends bursts in timeslots 1, 8, and 9, as indicated by edge T3 in the figure. Given a sufficiently long cycle time B, the source nodes will as a consequence of the

25   adaptive process described above, generally converge toward a state in which all are transmitting at or near the requested rates to all of their destinations.

When Destination receives a request, it checks the state of all the timeslots in its scheduling cycle. Timeslots that are unassigned and available for use belong to the AVAILABLE state, as indicated by reference numeral S1 in FIG. 5. In response to a

30   request for a number d of timeslots, Destination selects a number $\hat{d}$ from the AVAILABLE timeslots. Usually, it will be possible to fully satisfy the request, in

which case d and $\hat{d}$ will be equal. In some cases, $\hat{d}$ may be less than d, due, for example, to an insufficient number of AVAILABLE timeslots, or to a policy which restricts the maximum available rate. The state of the selected timeslots is changed from AVAILABLE to ASSIGNED, as indicated by edge E1 in FIG. 5. The

5 ASSIGNED state, indicated by reference numeral S2 in the figure, is given to those timeslots that have been assigned to a source but have not yet been used. The distinction between used and unused timeslots is important because, as noted above and as will be seen below, failure to use an assigned timeslot in the next cycle after it has been assigned is indicative of a conflict at the source.

10 Destination then responds to the request with a grant message identifying the $\hat{d}$ timeslots in which burst arrivals from Source are expected. Upon receipt of the grant message, Source computes the required departure time of each burst. If there is no source conflict, Source transmits $\hat{d}$ bursts in every cycle, beginning with the cycle after receipt of the grant message, until a need arises for Source to change its

15 transmission rate. In that event, Source sends a new request message.

When Destination receives the first burst in an assigned timeslot, it changes its state from ASSIGNED to USED, as indicated by edge E2 in the figure. The USED state is indicated in the figure by reference numeral S3.

As indicated in the Figure by edge E3, if Source makes a request to decrease

20 the rate or stop transmission, Destination will return some or all of the USED timeslots to the AVAILABLE state.

A timeslot is in the state CONFLICT, indicated by reference numeral S4 in the figure, if it is unavailable due to a source conflict. As noted, Destination will infer such a conflict when an assigned timeslot remains unused in a following cycle,

25 allowing time for a round-trip propagation delay. In that event, Destination will change the state of the timeslot to CONFLICT, as indicated by edge E4 of the figure.

It should be noted that if there is no conflict, but Source has no data to send in the timeslot, Source should send dummy bursts at the designated times to prevent Destination from making false detections of source conflicts.

30 According to one possible method of conflict resolution, Destination selects an AVAILABLE timeslot and assigns in place of the conflicted timeslot, meanwhile

changing the state of the newly assigned timeslot to ASSIGNED. The state of the
conflicted timeslot is concomitantly changed to AVAILABLE. These two state
changes are represented, respectively, by edges E5 and E6 of FIG. 5. The selection of
the substitute timeslot may, for example, be made randomly from those timeslots

5   which are AVAILABLE. The timeslot substitution is communicated by Destination
to Source in a "modify" message, upon receipt of which Source begins to use the new
timeslot if it is not subject to a conflict.

According to a more general approach to conflict resolution, the timeslot
substitution described above is the first of two alternative actions that Destination may

10   take, with respective probabilities p and $1-p$. The second alternative is to randomly
select a USED timeslot, thus bumping the current user of that timeslot. We refer to
such an action as "destination bumping."

If, for example, the selected timeslot is designated t_u, Destination will change
the state of t_u to ASSIGNED, and will send a modify message to the bumped source,

15   i.e., the source which had been using t_u, instructing it to begin using a new timeslot t_c.
If there is no conflict, the bumped source will then begin using timeslot t_c. The state
change of timeslot t_u is indicated in FIG. 5 by edge E7. Destination will of course
also send a modify message to Source, instructing it to begin using timeslot t_u.

Destination bumping tends to prolong the convergence time of the protocol.

20   However, numerical simulations suggest that by introducing a small amount of
destination bumping, e.g. by reducing p from 1.00 to about 0.98, it will in at least
some cases be possible to significantly reduce the minimum required cycle length
while taking only a modest penalty in convergence time.

Up to this point, our discussion has assumed that when there is a source

25   conflict, the only response available to Source is simply to refrain from transmitting
any burst that conflicts with one already scheduled. We will now describe an optional
further protocol which permits Source to schedule a new departure time for the
conflicting burst. To obtain the new departure time, Source takes it away from an
already-scheduled burst. For this reason, we refer to this further protocol as

30   "departure-time reassignment." We refer to the seizing of a departure time from an
already-scheduled burst as "source bumping."

We suppose, first, that destination j has sent Source a grant message containing an assigned timeslot. We suppose moreover that the burst-departure time to destination j conflicts with the burst-departure time to a different destination j'. In accordance with our protocol for departure-time reassignment, Source has two

5   alternative courses of action, with respective probabilities q and $1-q$. The first course of action is to stop transmitting bursts to destination j' and to begin transmitting bursts to destination j. The second alternative is simply to refrain from transmitting bursts to destination j.

In the case of either alternative, the respective destination will detect a source

10   conflict and will reassign the scheduled arrival time, as explained above.

To make the departure-time-reassignment protocol more robust, Source may optionally transmit a control burst to destination j' before stopping its burst transmission. Such a control burst will explicitly indicate that source bumping has occurred, and that destination j' must therefore schedule a new timeslot.

15   Source bumping affects the convergence time of the protocol for distributed scheduling. When q is near zero (nearly no source bumping), Source must rely entirely on each destination's conflict-detection and rescheduling operations to resolve conflicts. When q is near unity (source bumping almost always occurs when there is a conflict), many destinations may simultaneously be forced to reschedule

20   displaced bursts, resulting in still more bumping, which in the worst case may lead to divergence. Numerical simulations have suggested that a moderate amount of departure-time reassignment, e.g. q of about 0.3, can somewhat improve convergence times. Numerical simulations have also suggested that at least for synchronous traffic, the shortest cycle length sufficient to schedule all demands can be reduced

25   somewhat by departure-time reassignment.

Up to this point, our discussion has assumed that all traffic is in the same service class, and that all bursts have the same priority. In fact, however, our protocol can be modified to, e.g., prohibit bumping of specified classes of service, or to impose a hierarchy, i.e., to permit one service to bump another only if the bumped service has

30   lower or equal priority.

According to one approach which allows for priority service, the *destination*-bumping procedure described above is modified such that when Destination, with probability 1-*p*, makes a random selection of a USED timeslot, the random selection is limited to USED timeslots carrying services with priority lower than or equal to the

5   priority of the bumping service. The *source*-bumping procedure described above is also modified—in this case by introducing a new probability q', such that Source, with probability q', stops transmitting bursts to destination j' and starts transmitting bursts to destination j. The new probability q' is the same as the old probability q if and only if the service destined for j' has priority lower than or equal to the priority of

10   the service destined for j. Otherwise, $q'=0$.

All of the functions described above related to the scheduling of bursts at source and destination nodes are readily carried out by special or general purpose digital information processing devices acting under appropriate instructions embodied, e.g., in software, firmware, or hardware programming.

15   Other contexts to which the principles described above may be applied include, e.g., wireless local area networks (WLANs). In a WLAN, each burst is inherently labeled by a radiofrequency channel, such that different frequencies are received by different destination nodes. Control communications may be sent, e.g., on one or more dedicated control channels. Formats for radiofrequency bursts are

20   well known and need not be described in detail here.

Example

We performed numerical simulations of the blocking behavior of a network with distributed scheduling as described above, in which non-colliding timeslots are randomly assigned. We assumed that each of N source nodes requests d timeslots, so

25   the total number of requested timeslots is Nd. We considered the effect of a parameter ρ, referred to as the "utilization." The utilization is defined as the ratio of the total number of requested timeslots, to the cycle length B (as also measured in timeslots). FIG. 6 is a graph of the blocking probability as a function of the number of source nodes, when the destination nodes do not adapt to source conflicts. Curves

30   are provided at utilization levels of 0.43, 0.85, and 1.00. Substantial improvements in

performance are expected when there is adaptation to source conflicts, e.g. through one of the mechanisms for conflict resolution described above.